- 1 -

OBJECT TRACKING METHOD AND OBJECT TRACKING APPARATUS

CROSS-REFERENCE TO RELATED APPLICATION

The present invention is related to U.S. patent application Serial No. 09/592,996 filed June 13, 2000 entitled "OBJECT TRACKING METHOD AND OBJECT

5 TRACKING APPARATUS", the entire content of which is incorporated herein by reference for all purposes.

BACKGROUND OF THE INVENTION

The present invention relates to an object tracking system using an imaging unit, or in particular

10 to an object tracking method and an object tracking apparatus for automatically detecting an object within a monitor area by processing an image signal and tracking the object by controlling the universal head of the camera carrying the imaging unit.

15 An object tracking system such as a remote monitor system using an imaging unit (hereinafter referred to as "the camera") such as a TV camera has been widely used. Many of them are what is called a manned monitor system which is operated by a monitor

20 person watching the image displayed on the monitor. In the manned monitor system, however, a monitor person is required to keep watching the image displayed on the monitor and identify in real time an object such as a man or an automotive vehicle which may intrude into the

monitor area. This is a great burden on the monitor person.

The concentration power of the man is limited. In the manned monitor system, therefore, an intruding object may be unavoidably overlooked, thereby posing the problem of low reliability. Also, in the case where an intruding object is found in the image (camera image) picked up by the camera, the camera universal head (electrically operated swivel base) carrying the camera is required to be operated in such a manner as to catch the intruding object within the camera view filed (i.e. the imaging view field). Partly due to the recent explosive extension of ownership of the monitor camera, however, one monitor person often watches a multiplicity of camera images on a plurality of monitors. In the case where a plurality of cameras catch an intruding object at the same point time, the universal heads of the plurality of the cameras cannot be easily operated at the same time. In such a case, the intruding object is liable to be overlooked.

Demand is high, therefore, for what is called a monitor system of automatic detection and tracking type in which an intruding object is detected not by a monitor person but automatically by processing the image picked up by the camera (camera image), and controlling the universal head of the camera to catch the intruding object in the camera view field as

required thereby to make a predetermined announcement or alarm.

As a function to realize the monitor system described above, the camera universal head is required

5   to be controlled in such a manner that an object to be monitored and considered an intruding object is detected from an image signal by a predetermined monitor method and caught within the camera view field.

In the conventionally implemented monitor

10 system of this type, an intruding object in the view field is detected by the difference method. In the difference method, an image (input image) picked up by the imaging unit such as a camera is compared with an image containing no image of the object to be detected

15 (i.e. a reference background image prepared in advance), the difference of the brightness value is determined for each pixel or each pixel block including a plurality of pixels, and an area having a large difference (image signal change area) is detected as an

20 intruding object.

The image of the intruding object detected in this way is registered as a template, and the motion of the intruding object within the camera view field is detected thereby to control the camera universal head

25 in accordance with the motion of the object.

The process for this conventional method of detecting an intruding object is explained with reference to the flowchart of Fig. 5. In Fig. 5, the

process starts with the initialization step 101 in
which external devices, variables and an image memory
for executing the intruding object tracking method are
initialized. Next, in step 102 defined by dotted line,

5   an intruding object is detected by the difference
method. Step 102 includes a first image input step
102a for acquiring an input image having 320 row pixels
and 240 column pixels, for example, from the camera.
Next, in the difference processing step 102b, the

10 brightness value difference for each pixel between the
input image acquired in the first image input step 102a
and the reference background image prepared in advance
is calculated as a brightness value of the difference
image. Then, in the binarization step 102c, a binary

15 image is acquired from the pixel value of the
difference image (difference value) obtained in the
difference processing step 102b in such a manner that
the value of each pixel less than a predetermined
threshold value Th (say, 20) is regarded as "0" and the

20 pixel value not less than the threshold value Th as
"255" (calculated with each pixel as 8 bits in pixel
value). Next, in the labeling step 102d, a cluster of
pixels having the value "255" is detected in the binary
image obtained in the binarization step 102c, and each

25 pixel is discriminated by a number attached thereto.
In the intruding object presence determining step 102e,
it is determined that an intruding object exists in the
monitor area in the case where the cluster of the

pixels having the value "255" numbered in the labeling step 102d meets predetermined conditions (for example, a size having 20 row pixels and 50 column pixels) as described above with reference to the difference

5 method.

In the case where it is determined that an intruding object exists in the intruding object determining step 102e, the process branches to the first alarm/alarm display step 103. In the case where

10 it is determined that there is no intruding object, on the other hand, the process branches to the first image input step 102a again thereby to execute the difference method.

Next, in the first alarm/alarm display step

15 103, the monitor person is informed, for example, by sounding an alarm indicating that an intruding object has been found or displaying an alarm indicating that an intruding object has been found on the monitor. Then, in the template registration step 104, an image

20 of the intruding object is cut out from the input image based on the cluster of the pixels having the pixel value "255" numbered in the labeling step 102d, and registered as a template.

Next, the position is detected where the

25 degree of coincidence (likeliness) becomes maximum between an image sequentially input and the template thereby to detect the position of the intruding object. This method is widely known as the template matching,

and is described in detail, for example, in "Digital
Picture Processing" published by ACADEMIC PRESS pp.
296-303, 1976 and U.S. Patent No. 5,554,983, the
disclosure of which is hereby incorporated by reference
5   herein.

The template matching is used by reason of
the fact that the execution of the difference method
requires the reference background image 602, and in the
case where the camera universal head is controlled in
10   such a manner as to catch an intruding object within
the camera view field, the optical axis of the camera
deviates undesirably, thereby making it impossible to
use the reference background image 602 prepared in
advance.

15   Normally, in the case where the position of
an intruding object is detected using the template
matching, the position change of the intruding object
is followed in such a manner that the image at the
position of the intruding object detected by template
20   matching is sequentially updated as a new template.
This process is executed in and subsequent to the
second image input step 105, and explained below.

In the second image input step 105, like in
the first image input step 102a, an input image having
25   320 row pixels and 240 column pixels, for example, is
acquired from the camera. Next, in the template
matching step 106, an image having the highest degree
of coincidence with the template is detected from the

input image acquired in the second image input step

105.   Normally, the job of comparing the whole input

image with the template requires a long calculation

time.   Therefore, a predetermined range with respect to

5   the template (for example, a range extended 20 column

pixels and 50 row pixels with respect to the template)

is set as a search area, and an image highest in the

degree of coincidence with the template is detected

within this search area.

10   The degree of coincidence can be calculated

using the normalized cross-correlation value $r(\Delta x, \Delta y)$

described in US Patent Serial No. 09592,996.

        The normalized cross-correlation value $r(\Delta x, \Delta y)$ is included in the range $-1 \leq r(\Delta x, \Delta y) \leq 1$, and in

15   the case where the input image and the template

coincide completely with each other, the value of "1"

is assumed.   The template matching is a process in

which Δx and Δy are scanned within the search range,

that is to say, changed in the range $-50 \leq \Delta x \leq 50$ and

20   $-20 \leq \Delta y \leq 20$, respectively, thereby to detect the

position associated with the maximum normalized cross-

correlation value $r(\Delta x, \Delta y)$.

        Next, in the coincidence degree determining

step 107, the degree of coincidence $r(\Delta x, \Delta y)$ is

25   determined.   In the case where the normalized cross-

correlation value is used, and not less than 0.7, for

example, it is determined that the degree of

coincidence is high, and the process branches to the

intruding object position correcting step 108. In the case where the normalized cross-correlation value is less than 0.7, on the other hand, it is determined that the degree of coincidence is low, and the process

5 branches to the first image input step 102a.

A high degree of coincidence is indicative of the presence of an image similar to the template in the input image, i.e. the presence of an intruding object in the input image. In this case, the intruding object

10 continues to be tracked.

A low degree of coincidence, on the other hand, indicates the absence of an image similar to the template in the input image, i.e. the absence of an intruding object in the input image. In this case, the

15 process branches to the first input image input step 102a, and the process for detecting an intruding object is executed again by the difference method.

In the intruding object position correcting step 108 executed in the case where the degree of

20 coincidence is high, the value $(x0+\Delta x, y0+\Delta y)$ is corrected as a new position of the intruding object based on the position $(\Delta x, \Delta y)$ associated with the maximum degree of coincidence. Next, in the template update step 117, the input image obtained in the second

25 image input step 105 is cut out as a new template image based on the newly determined position of the intruding object.

Further, in the camera universal head control

step 118, the camera universal head (i.e. the direction of the optical axis of the camera) is controlled according to the displacement between the position of the intruding object detected in the template matching step 106 and a predetermined reference position of the input image (i.e. a predetermined reference position in the imaging view field) such as the center of the input image. As an example, assume that an intruding object is detected at a position 802 shown in Fig. 6. Assuming that the center position of the intruding object coincides with the center 803 of the template, the displacement dx, dy from the center of the image is calculated.

In the case where the template center position 803 is located leftward of the center 804 of the input image by at least a predetermined amount s (dx < -s), the camera universal head is panned leftward. In the case where the template center position 803 is located rightward of the center 804 of the input image by at least a predetermined amount s (dx > s), on the other hand, the camera universal head is panned rightward. In the case where the intruding object is located at about the center of the image ($-s \leq dx \leq s$), the camera universal head is not required to be controlled. Therefore, the position where the camera universal head begins to be controlled can be designated by the predetermined amount s. The predetermined amount s is 50, for example.

Also, in the case where the center position 803 of the template is higher than the center 804 of the input image (dy < -s), the camera universal head is tilted upward, while in the case where the center position 803 of the template is lower than the center 804 of the input image (dy > s), the camera universal head is tilted downward.

As an alternative, the control speed of the pan motor and the tilt motor may be changed according to the absolute value of dx and dy (the control speed is higher, the larger the absolute value of dx or dy).

Finally, in the second alarm/alarm display step 119, an alarm is sounded, for example, to inform the monitor person that an intruding object is being tracked. Alternatively, an alarm indicating that an intruding object is being tracked may be displayed on the monitor.

The method of tracking an intruding object using the template matching described above poses the problem that in the case where the direction of the intruding object to be tracked changes (in the case where the intruding person turns his head to the right or looks back, for example), the displacement between the intruding object and the position detected by the template matching is increased and the accurate and stable tracking becomes impossible.

This is due to the characteristic of the template matching in which a pattern having a high

contrast in the template (the part of a large
brightness value difference) is matched. Assume that
an intruding object is an automotive vehicle which is
first directed forward and substantially wholly held in
5 the template. Once the vehicle changes its running
direction sideways, for example, other parts than the
front part of the vehicle can no longer be held in the
template. Unlike during the time when the whole
vehicle is held in the template, the center of the
10 template moves to the front part from the center of the
vehicle, and therefore the detection position of the
intruding object is displaced.

This phenomenon is explained with reference
to Fig. 7. In order to simplify the explanation, this
15 figure shows an example where the camera universal head
is not controlled. To explain the phenomenon in which
an intruding object becomes impossible to hold in the
template by the template matching method of tracking
the intruding object, the process executed in the case
20 where a vehicle running along a curved road in the
imaging view field is tracked as an intruding object is
shown in Fig. 7.

Reference numerals 901, 903, 905 and 907
designate template images at time points t_1-1, t_1, t_1+1
25 and t_1+2, respectively, and numerals 901a, 903a, 905a
and 907a the template at time points t_1-1, t_1, t_1+1 and
t_1+2, respectively. Numerals 902, 904, 906 and 908
designate input images at time points t_1, t_1+1, t_1+2 and

t_1+3, respectively, and numerals 902a, 904a, 906a and 908a the template positions at time points t_1-1, t_1, t_1+1 and t_1+2, respectively (the positions of the intruding object at time points t_1-1, t_1, t_1+1 and t_1+2,

5   respectively). Also, numerals 902b, 904b, 906b and 908b designate the positions of the intruding object detected by template matching at time points t_1, t_1+1, t_1+2 and t_1+3, respectively.

SUMMARY OF THE INVENTION

10   In Fig. 7, the template 901a registered at time point t_1-1 is an image in which the vehicle the front portion of the vehicle is directed substantially forward. At time point t_1, the template matching is conducted using the template 901a (the template

15   matching step 709) thereby to detect the position to which the intruding object has moved, while at the same point time updating the template 901a to the template 903a (the template update step 710).

       Then, the template is updated to the template

20   905a at time point t_1+1 and to the template 907a at time point t_1+2. After this process is repeated up to time point t_1+3, the template, that held the front portion of the vehicle with the head lights when the tracking started at time point t_1-1, now holds the left side of

25   the vehicle at time point t_1+3. This phenomenon is caused by the fact that the matching is conducted in such a manner as to reduce the displacement between the

input image handled by template matching and the high-contrast portion of the template image, i.e. the head lights of the vehicle in this case. As a result, as shown in Fig. 7, the template is displaced leftward

5 when the intruding object turns from right to left, and rightward when it turns from left to right.

Further, although the template 901a which held only the image of the vehicle at time point t_1-1, the template 908b at time point t_1+3 contains a

10 background image other than the intruding object to be tracked as the result of displacement with respect to the intruding object. In the case where the template matching process is continued using the template 908b containing many images other than the intruding object

15 to be tracked, the intruding object cannot be caught but the background portion contained in the template is matched. Thus, the method of tracking an intruding object using the template matching has the disadvantage that in the case where the intruding object changes its

20 direction, the pattern of the intruding object apparently moves and displaces the template position from the position of the intruding object. As a result, the accurate tracking of the intruding object to be tracked cannot be guaranteed and the steady

25 tracking of the intruding object becomes impossible.

As described above, the prior art has the disadvantage that steady tracking becomes impossible in the case where an intruding object changes its

direction considerably.

An object of this invention is to provide an object tracking method and an object tracking apparatus free of the disadvantages of the prior art.

5   Another object of the invention is to provide an object tracking method and an object tracking apparatus which operates reliably and stably to detect and track an intruding object accurately even in the case where the intruding object changes its direction

10 considerably.

In order to achieve the objects described above, according to one aspect of the invention, there is provided an object tracking method for detecting and tracking an object in a picked-up image based on an

15 image signal acquired by an imaging unit, including the steps of:

detecting an image change area between at least two frames of images picked up at different time points by the imaging unit; and

20   correcting the object position by detecting the position of the object based on the detected image change area and setting the detected object position as a new detected object position in place of the present detected object position.

25   In an example of this aspect of the invention, the image change area detection step is to detect a change area of the image signal by comparing the two frames of image signals picked up at different

time points by the imaging unit while the camera

universal head (imaging unit controller) is stationary.

In another example of this aspect of the invention, the

object position correcting step is to enlarge the

5 intended area in steps for accurate position detection

of an intruding object based on the intruding object

position information and to correct the intruding

object position information by detecting the change

area of the image signal obtained in the image signal

10 change area detection step and the accurate position of

the intruding object based on the intended area.

In order to achieve the objects of the

invention, according to another aspect of the

invention, there is provided an object tracking

15 apparatus for detecting and tracking an object in a

picked-up image based on an image signal acquired by

an imaging unit, incuding:

an image input unit for converting the video

signals acquired by the imaging unit sequentially into

20 image signals; and

a processing unit for processing the image

signals converted by the image input unit, in a

predetermined sequence;

wherein the processing unit detects an image

25 change area between at least two frames of images

acquired at different time points from the imaging

unit, and by detecting the position of the intruding

object based on the detected image change area, sets

the detected object position as a new detected object
position in place of the present detected object
position.

According to this invention, the employment
5 of the inter-frame difference processing step obviates
the problems of the prior art in which the template may
be displaced with respect to an intruding object and
come to contain many images other than the intruding
object in the case where the intruding object changes
10 its direction, with the result that the template
matches with the background contained therein and the
image of the intruding object cannot be caught.
Further, even an intruding object considerably changing
its direction can be detected and tracked accurately.

15 Other objects, features and advantages of the
invention will become apparent from the following
description of the embodiments of the invention taken
in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

20 Figs. 1A and 1B are flowcharts for explaining
an example of the processing flow according to a third
embodiment of the invention.

Fig. 2 is a flowchart for explaining an
example of the processing flow according to a first
25 embodiment of the invention.

Fig. 3 is a flowchart for explaining an
example of the processing flow according to a second

embodiment of the invention.

Fig. 4 is a diagram showing an example of the configuration of a monitor system embodying the invention.

5   Fig. 5 is a flowchart for explaining the processing flow according to the conventional object tracking method.

Fig. 6 is a diagram for explaining a method of controlling the universal head of the camera.

10   Fig. 7 is a diagram for explaining the problems of the conventional object tracking method.

Figs. 8A to 8E are diagrams for explaining a method of detecting the object position by the inter-frame difference method according to this second

15 embodiment of the invention.

Fig. 9 is a diagram for explaining an example of the processing flow in an object tracking method according to the second embodiment of the invention.

Fig. 10 is a diagram for explaining a method

20 of detecting the object position according to the third embodiment of the invention.

Figs. 11A and 11B are diagrams for explaining the effects of the second embodiment of the invention.

Figs. 12A to 12E are diagrams for explaining

25 a method of detecting the object position by the inter-frame difference method according to this first embodiment of the invention.

Figs. 13A to 13C are timing charts showing

the operations of a template matching processing, a
camera universal head control processing and an
intruding object position correcting processing using
an inter-frame difference processing, according to the

5 invention.

DETAILED DESCRIPTION OF THE INVENTION

An object tracking method and an object
tracking apparatus according to embodiments of this
invention are explained in detail below with reference

10 to the accompanying drawings.

First, Fig. 4 shows an example of the
hardware configuration of a monitor system according to
an embodiment of the invention. Reference numeral 401
designates an imaging unit including at least a camera

15 401a, a zoom lens 401b and a universal head
(electrically-operated swivel base) 401c. Numeral 402
designates a processing unit including an image input
unit (image input I/F) 402a, a universal head control
unit (universal head input I/F) 402b, a lens control

20 unit (lens control I/F) 402c, an operation input unit
(operation input I/F) 402d, an image memory 402e, a MPU
(microprocessing unit) 402f, a work memory 402g, an
external input/output unit (external input/output I/F)
402h, an image output unit (image output I/F) 402i, an

25 alarm output unit (alarm output I/F) 402j and a data
bus 402k. These component parts are combined with an
operating unit 403, an external storage unit 404, an

image monitor 405 and an alarm lamp 406.

The operating unit 403 includes a joystick 403a, a first button 403b and a second button 404c. The output terminal of the camera 401a is connected to

5   the data bus 402k through the image input unit 402a. The control unit of the zoom lens 401b is connected to the data bus 402k through the lens control unit 402c. The universal head on which the camera 401a is mounted is connected to the data bus 402k through the universal

10 head control unit 402b. The output terminal of the operating unit 403 is connected to the data bus 402k through the operation input unit 402d.

The external storage unit 404 is connected to the data bus 402k through the external input/output

15 unit 402h. The image monitor 405 is connected to the data bus 402k through the image output unit 402i. The alarm lamp 406 is connected to the data bus 402k through the alarm output unit 402j. The MPU 402f and the work memory 402g are connected directly to the data

20 bus 402k.

The camera 401a catches the monitor area within a predetermined view field and by picking up an image of the monitor area, outputs a video signal. For this purpose, the camera 401a has a zoom lens 401b and

25 is mounted on the universal head 401c. The video signal picked up is stored in the image memory 402e from the image input unit 402a through the data bus 402k. The external storage unit 404 functions to store

the programs and data, which are read into the work memory 402g through the external input/output unit 402h on the one hand and held in the image storage unit 404 on the other hand, as required.

5   The MPU 402f executes the process according to the program held in the external storage unit 404 and read into the work memory 402g when the processing unit 402 is in operation, and analyzes the images stored in the image memory 402e in the work memory

10   402g.   In accordance with the result of this processing, the MPU 402f controls the zoom lens 401b and the universal head 401c thereby to change the imaging view field of the camera 401a.   At the same time, MPU 402f displays an image of the result of

15   detection of an intruding object on the monitor 405 and turns on the alarm lamp 406 as required.

          Next, an example of the operation according to an embodiment of the invention is explained.   The embodiments explained hereinafter are operated with the

20   hardware configuration shown in Fig. 4.   First, the operation of the first embodiment of the invention is explained with reference to Fig. 2.   Fig. 2 is a flowchart showing an example of the processing according to the first embodiment of the invention.

25   The process shown in Fig. 2 starts with the initialization step 101.   In the initialization step 101, the image memory area and the recording area are initialized.   Also, the reference background image

recorded in the external storage unit 404 is read and stored in the image memory 402e.

The image memory area is used to execute the process of tracking an intruding object, and secured in the image memory 402e. The recording area, on the other hand, is an area secured in the work memory 402g to store the variables, etc. required for executing the process.

In the intruding object detection processing step 102, the same process is executed as in the intruding object detection processing step according to the prior art described above with reference to Fig. 5. In Fig. 2, the respective steps making up the intruding object detection processing step 102 are indicated as a single step. The operation of each step is similar to the processing operation of the corresponding step in the prior art described above and therefore is not described again.

In the first alarm/alarm display step 103, an image of an intruding object detected is displayed on the image monitor 405, for example, and the alarm lamp 406 is turned on to announce that an intruding object exists in the monitor area.

In the image input step 105, the image signal picked up by the camera 401a is input through the image input unit 402a thereby to acquire an input image having 320 row pixels and 240 column pixels, for example, with eight data bits per pixel (320 × 240

pixels, 8 bits/pixel). In the process, the input image acquired in the preceding processing frame (in the step one session before) is stored in the image memory 402e. The input image stored in the image memory 402e is

5 erased upon the lapse of a predetermined length of time (say, ten seconds) after the input time.

In the camera universal head stop determining step 109, it is determined whether the universal head 401c is out of operation, i.e. whether the imaging

10 direction (that is to say, the direction of the optical axis) of the camera are fixed or not. In the case where the universal head 401 is out of operation, the process branches to the inter-frame difference processing step 110. Otherwise, the process branches

15 to the universal head control step 118 (for the reason described later).

In the inter-frame difference processing step 110, the processing of the inter-frame difference method is executed. According to the inter-frame

20 difference method, the input image of the preceding processing frame stored in the image memory 402e is used in place of the reference background image 602 in the conventional difference method for detecting an intruding object described in Fig. 6. Specifically, in

25 the case where two frames of input images obtained at different time points are used by the difference method, the portion moving in the input image (the portion with the pixel value undergoing a change) is

detected. Since the intruding object is in motion, the pixel values of the intruding object develop a difference thereby making it possible to detect the intruding object. This difference method is

5   specifically called the inter-frame difference method and widely used. Incidentally, the method of determining the difference between the input image and the reference background image described with reference to Fig. 6 is called the background difference method.

10   Next, in the maximum difference amount position search step 112, the difference image obtained in the inter-frame difference processing step 110 is searched for an area containing many differences. This processing is explained with reference to Figs. 12A to

15   12E. Fig. 12A shows an image 1001 input in the image input step 105, and Fig. 12B an input image 1002 stored in the image memory 402e. The input image 1002 is acquired at a time point different from the input image 1001 by a predetermined length of time (say, one

20   second). Then, a hatched area 1003 in Fig. 12C is obtained as a difference image.

An area expanded from the intruding object position 1003a (at this time point, the position of the intruding object in the preceding processing frame is

25   used as the present position of the intruding object is unknown) along Y and Y directions by a predetermined number d of pixels (say, 100) is set as a search area 1003b. The accumulated value of the pixel difference

in this area is projected on X and Y axes. Projected
images 1003c and 1003d shown in Fig. 12C are obtained.
These images are denoted as hx(x) and hy(y),
respectively. In Fig. 12C, the graph projected is

5   shown in superposition on the difference image to
facilitate the understanding.

Fig. 12D shows hx(x) representing the
projection on the X axis of the accumulated value of
the differences of the images in the search area. The

10 area 1003a indicated by dotted line represents a range
of the detection position of the intruding object in
the preceding processing frame. The range associated
with the maximum accumulated value is detected within
the width tx of the intruding object (the width of the

15 intruding object detected in the intruding object
detection step 102). In this example, the area 1003f
defined by solid line is the particular range. This is
expressed by the equation below.

$$x_1 = \mathop{arg\,max}\limits_{x_0 - d < x_1 < x_0 + d} \sum_{i=0}^{t_x-1} h_x(x_1 + i) \qquad (1)$$

In equation (1), x_0 designates the position of

20 the intruding object (with respect to the coordinate of
the upper left point of the rectangle circumscribing
the intruding object) detected in the preceding frame,
and x_1 designates the position associated with the
greatest difference within the width tx of the

25 intruding object detected. A moving object exists in

the area where the greatest difference is detected. In similar fashion, in the graph shown in Fig. 12E with the difference projected on Y axis, an area 1003h with the greatest difference detected within the height ty

5 of the intruding object (the height of the intruding object detected in the intruding object detection processing step 102) is detected from the area 1003g where the intruding object exists in the preceding processing frame indicated by dotted line.

10   This is expressed by equation (2) below.

$$y_1 = \mathop{\textbf{\textit{arg max}}}_{y_0 - d < y_1 < y_0 + d} \sum_{j=0}^{t_y - 1} h_y(y_1 + j) \qquad (2)$$

In equation (2), y_0 designates the position of the intruding object detected in the preceding frame, and y_1 the position associated with the greatest difference within the height ty of the intruding object

15 detected.

In the inter-frame difference processing step 110, therefore, the position (x_1, y_1) associated many differences is detected.

Incidentally, according to this embodiment,

20 as shown by equations (1) and (2), the values x_1, y_1 associated with the largest accumulated value of $hx(x)$, $hy(y)$ are determined in the range of $x_1 < x < x_1+dx$, $y_1 < y < y_1+dy$ where x_1 or y_1 is changed in the range $x_0-d < x_1 < x_0+d$ or $y_0-d < y_1 < y_0+d$, respectively. As an

25 alternative, in the case where $\sum_{i=0}^{dx} hx(x_1 + i)$, $\sum_{i=0}^{dy} hy(y_1 + i)$

in equations (1) and (2) exceed a predetermined
threshold value while x_1 and y_1 change in the range of
$x_0 - d < x_1 < x_0 + d$, $y_0 - d < y_1 < y_0 + d$, the calculation of
equations (1) and (2) is stopped, and the prevailing x_1

5   or y_1 may be used as the corrected position of the
template. In this case, the predetermined threshold
value is set to 30 % of the maximum accumulated value
$255 \times (dy + 2d)$ ($255 \times (dx + 2d)$ for y axis), for
example. The range detected by this value represents a

10   portion including the difference of not less than 30 %
of the maximum accumulated value of the difference,
thereby making it possible to reduce the amount of
calculation of equations (1) and (2).

        According to the above-described method, the

15   difference obtained by the inter-frame difference
processing is projected on the X and Y axes, then the
area having the maximum difference is obtained based on
the differences thus projected, and the position of the
intruding object is corrected Instead of this method,

20   another method may be employed in which an area of a
template size within the search range is scanned from
the uppermost left side to the lowest right side, for
example, within the camera view field and a portion
where the total amount of the difference contained

25   within the template size area becomes maximum is
detected thereby to correct the position thus detected
as a new position of the intruding object. This method
may be modified in a manner that the scanning within

the search range is stopped when the total amount of the difference within the template size area becomes equal to or larger than a predetermined value thereby to correct the area thus detected as a new position of

5 the intruding object.

Next, in the difference amount determining step 201, the total value of the differences in the area $(x_1, y_1) - (x_1+tx, y_1+ty)$ containing a large difference is determined. In the case where this value

10 is less than $tx \times ty \times th \times 0.1$, the difference amount is small, i.e. the object moving in the detected area is determined not as an intruding object, and the process branches to the intruding object detection processing step 102. In the case where the value is

15 not less than $tx \times ty \times th \times 0.1$, on the other hand, the difference amount is large, i.e. the object moving in the detected area is determined as an intruding object, and the process branches to the intruding object position correcting step 115. The value 0.1, for

20 example, is to determine whether the difference corresponding to ten percent, for example, of the threshold value is included or not in the detected area. The value 0.1 (ten percent) is an example after all, and the ratio can be set arbitrarily in accordance

25 with the actual setting of the system.

In the intruding object position correcting step 115, the detected position (x_1, y_1) is replaced by the position (x_0, y_0) where the intruding object exists.

Only in the case where it is determined in
the universal head stop determining step 109 that the
camera universal head 401c is stopped, the intruding
object position correcting step 115 is executed from
5   the inter-frame difference processing step 110. This
is due to the fact that the inter-frame difference
method is executed by comparing the input images of two
frames obtained at different time points, and therefore
the inter-frame difference processing cannot be
10   executed while the camera universal head 401c is
moving.

Next, in the camera universal head control
step 118, the universal head of the camera is
controlled in the same manner as the conventional
15   camera universal head control method explained above
with reference to Figs. 5 and 6. Further in the second
alarm/alarm display step 119, the image of an intruding
object being tracked is displayed on the image monitor
405 or the alarm lamp 406 is turned on thereby to
20   inform that the intruding object is being tracked in
the monitor area.

Specifically, according to the embodiment
shown in Fig. 2, the problem can be solved that in the
case where an intruding object in motion exists in the
25   detected area and the template matching is continued
using a template containing many images other than the
intruding object to be tracked, by the intruding object
tracking method using the template matching, the

intruding object could not be grasped but the
background portion included in the template is matched,
so that the tracking of the intruding object to be
tracked cannot be guaranteed and the steady tracking is
5 made impossible. In this way, the intruding object in
motion can be positively tracked. As a result, the
performance of the intruding object tracking process is
improved to facilitate the construction of a highly
reliable monitor system.

10 The operation according to the second
embodiment of the invention is explained with reference
to Fig. 3. Fig. 3 is a flowchart showing an example of
the processing in the second embodiment of the
invention.

15 The processing shown in Fig. 3 is a
combination of the processing of Fig. 2 and the
conventional intruding object tracking method using the
template matching shown in Fig. 5. In this combined
processing, the steps other than the difference amount
20 determining step 113 and the intruding object stop
determining step 116 are already explained and
therefore not explained below. Since the maximum
difference amount position search step 112 in this
embodiment is slightly different from that of the first
25 embodiment shown in Fig. 2, the maximum difference
amount position search step 112 in this embodiment will
be explained later with reference to Figs. 8A to 8E.

In this process, the position of the

intruding object is detected by template matching
through the process from the template matching step 106
to the first intruding object position correcting step
108. In view of the likelihood of the detected

5 intruding object position being displaced from the
template position as in the prior art, however, the
intruding object is corrected to the right position
through the following process of the camera universal
head stop determining step 109 to the second intruding

10 object position correcting step 115.

In the difference amount determining step 201
according to the first embodiment, the tracking of an
intruding object is stopped in the case where the
detected position difference amount is not more than a

15 predetermined value, and the process returns to the
intruding object detection processing step 102 again.
In the difference amount determining step 113 according
to the third embodiment shown in Fig. 3, however, the
tracking of the intruding object is continued even in

20 the case where the difference amount is not more than a
predetermined value.

In the case where the difference amount is
not more than a predetermined value, however, the
position of the intruding object is not corrected and

25 the process branches to the intruding object stop
determining step 116.

From the intruding object stop determining
step 116, the process branches to the intruding object

detection processing step 102 in the case where the detected position of the intruding object remains unchanged for a predetermined length of time (say, ten seconds), while the process branches to the template

5 update step 117 in the case where the detected position of the intruding object is found to have changed.

According to the first embodiment of the invention described above, the position of the intruding object is detected using only the inter-frame

10 difference method. In the inter-frame difference method, however, the difference ceases to appear in the case where the motion of the intruding object stops, in which case the tracking of the intruding object is immediately stopped and a new intruding object is

15 looked for. According to the second embodiment of the invention, on the other hand, the tracking of the intruding object can be continued using the template matching before the lapse of a predetermined time (say, ten seconds in the aforementioned case) even after the

20 motion of the intruding object stops.

The flow of this processing is explained with reference to Fig. 9. Fig. 9 is a diagram showing an example of the process executed according to the second embodiment of the invention and represents a case in

25 which a vehicle running along a curved vehicle lane in the imaging field of view shown in Fig. 7 is tracked as an intruding object.

Fig. 9 shows a case where the camera

universal head is not controlled in order to explain the effects of the inter-frame difference processing step 110 and the maximum difference amount position search step 112. The operation timings of the camera

5 universal head control processing and the inter-frame difference processing will be explained later.

Numerals 1101, 1103, 1105 and 1107 designate the template image at time points t_1-1, t_1, t_1+1 and t_1+2, respectively, and numerals 1101a, 1103a, 1105a and

10 1107a the template at time points t_1-1, t_1, t_1+1 and t_1+2, respectively. Also, numerals 1102, 1104, 1106 and 1108 designate input images at time points t_1, t_1+1 and t_1+2 and t_1+3, respectively, numerals 1102b, 1104b, 1106b and 1108b the template position at time points

15 t_1-1, t_1, t_1+1 and t_1+2, respectively (the position of the intruding object at time points t_1-1, t_1, t_1+1 and t_1+2, respectively), and numerals 1102a, 1104a, 1106a and 1108a the position of the intruding object detected by template matching at time points t_1, t_1+1 and t_1+2 and

20 t_1+3, respectively.

In Fig. 9, the template 1101a registered at time point t_1-1 is an image of the vehicle with the front portion thereof facing substantially forward. At time point t_1, the template matching is conducted using

25 the template 1101a (template matching step 709) to detect the moved position of the intruding object. At the same time, the position of the intruding object is accurately corrected by the inter-frame difference

method (the inter-frame difference processing step 110
to the intruding object correcting step 115). Further,
the template 1101a is updated to the template 1103a
(template update processing step 710).

5   Then, the template 110a is updated to the
template 1105a at time point t_1+1, and to the template
1107a at time point t_1+2. This process is repeated up
to time point t_1+3. Thus, the template held the front
portion with the head lamps of the vehicle at the

10   tracking start time point t_1. At time point t_1+3,
however, the template is displaced to the left side of
the vehicle according to the prior art shown in Fig. 7
(template 907a). According to the second embodiment of
the invention, in contrast, the template accurately

15   holds the intruding object (template 1107a).

         Even in the case where the intruding object
changes in direction as shown in Fig. 9, therefore, the
displacement between the intruding object and the
template which occurs with the direction change of the

20   intruding object in the conventional template matching
method is obviated, and the intruding object can be
tracked accurately.

         According to this embodiment, the position
(range) of the intruding object is specified by the

25   template matching step 106, then the search range is
set based on this specified range, and the inter-frame
difference processing and the maximum difference amount
position search processing are performed using the

search range. As compared with the range according to

the first embodiment of the invention (d = 100 in the

aforementioned example), therefore, the range can be

reduced to d = 50 pixels, for example, and therefore the

5   calculation amount in the inter-frame difference method

can be reduced.

The processing of the maximum difference

amount position search step 112 in the second

embodiment will be shown in Figs. 8A to 8E. In these

10   figures, the explanation of like portions as those of

Figs. 12A to 12E will be omitted. As shown in Fig. 8C,

an area which is made larger by a predetermined number

of pixels d, for example, 50 pixels in each of the X

and Y axes with respect to the intruding object

15   position 1003a is set as a search area 1003b.

Further, according to the second embodiment,

an intruding object can be tracked more accurately than

in the first embodiment in the case where a plurality

of objects exist in the imaging view field of the TV

20   camera. This advantage is explained with reference to

Figs. 11A and 11B. Fig. 11A shows an example in which

another object 1301b appears before the vehicle 1301a

tracked in the imaging view field. The object 1301b

which is located before the vehicle 1301a has a size

25   and a speed apparently larger than the vehicle 1301a.

In this case, the difference obtained by the inter-

frame difference processing is larger for the object

1301b. Assume that the detection position of the

intruding object in the preceding frame is indicated by the range of 1301c. According to the first embodiment of the invention, the range of processing the inter-frame difference is indicated by 1301d. The object 1301b is displayed in the range 1301d. In the case where the maximum difference position search step 112 is executed under this condition, the range in which the resulting accumulated difference value is maximum is represented by the object 1301b.

According to the second embodiment, in contrast, as shown in Fig. 11B, the range 1302a where the intruding object exists is roughly detected in the template matching step 106, so that the range 1302b where the inter-frame difference is processed is determined based on the range 1302a. In the case under consideration, the object 1301b is not displayed in the range 1302b, and even in the case where the maximum difference position search step 112 is executed, the range of the maximum accumulated difference value is not represented by the object 1301b unlike in the first embodiment, thereby making it possible to grasp the vehicle 1301a accurately.

As a result, the performance of the intruding object tracking process is improved and a reliable monitor system is easily constructed.

Figs. 13A to 13C are timing charts showing the operations of the template matching processing, the camera universal head control processing and the

intruding object position correcting processing using the inter-frame difference processing. Figs. 13A to 13C are diagrams showing the operations of the respective processings during a time period from a time point t0 to a time point t0+12. In the figures, a reference numeral 1401 depicts an axis representing the flow of the time points, in which the time points t0, t0+1, --- are marked sequentially in accordance with the processing cycles of the template matching processing. A reference numeral 1402 of Fig. 13A depicts the operation of the template matching processing, wherein each of time zones (for example, 1402a, 1402b, --) shown by hatched areas represents a time period during which the processing of the template matching step 106 is actually performed.

A reference numeral 1403 of Fig. 13B depicts the operation of the camera universal head control step 118, wherein each of time zones shown by hatched areas represents a time period during which the camera universal head is actually controlled. The time zones during which the camera universal head is actually controlled will be explained sequentially. If, in the processing frame at the time point t0, the detected position of the intruding object is deviated from almost the center portion of the camera view field, the control operation of the camera universal head is started at the camera universal head control step 118. When, in the processing frame at the time point t0+1,

the detected position of the intruding object is held

almost at the center portion of the camera view field,

the control operation of the camera universal head is

stopped at the camera universal head control step 118.

5 That is, the control time period of the camera

universal head is shown by a time period 1403a.

Similarly, when, in each of the processing frames at

the time points t0+3, t0+4 and t0+5, the detected

position of the intruding object is deviated from

10 almost the center portion of the camera view field and

then the detected position of the intruding object is

held almost at the center portion of the camera view

field in the processing frame at the time point t0+6,

the camera universal head is controlled during a time

15 period 1403b. Further, when, in each of the processing

frames at the time points t0+9 and t0+10, the detected

position of the intruding object is deviated from

almost the center portion of the camera view field and

then the position of the intruding object detected in

20 the processing frame at the time point t0+11 is held

almost at the center portion of the camera view field,

the camera universal head is controlled during a time

period 1403c.

A reference numeral 1404 of Fig. 13C depicts

25 the operation of the intruding object position

correction processing (steps 110 to 115) based on the

difference obtained in the inter-frame difference

processing step 110, wherein each of time zones shown

by hatched areas represents a time period during which
the inter-frame difference processing step 110 etc. are
executed. As described above, in the processing frame
at the time point t0+1, since the detected position of
5   the intruding object is held almost at the center
portion of the camera view field, the control operation
of the camera universal head is stopped at the end
point of the time period 1403a. Then, at the
processing frame of the next time point t0+2, since it
10 is determined that the control operation of the camera
universal head is stopped at the camera universal head
stop determining step 109, the inter-frame difference
processing is executed. That is, during the time
period 1404a, the inter-frame difference processing
15 step 110 is executed thereby to correct the position of
the intruding object (steps 112 to 115). At the time
period 1404b of the processing frame of the succeeding
time point t0+3, although the inter-frame difference
processing is also executed, the position of the
20 intruding object detected at the processing frame of
the time point t0+3 is deviated from almost the center
portion of the camera view field. Thus, the inter-
frame difference processing is not executed in the
processing frame of the succeeding time point t0+4.
25 Similarly, in the processing frame at the time point
t0+6 where the detected position of the intruding
object is held almost at the center portion of the
camera view field, the control operation of the camera

universal head is stopped (a time period 1403b). Then, the inter-frame difference processing step 110 etc. are executed in the processing frames of time points t0+7, t0+8 and t0+9 until the control operation of the camera

5   universal head is started again, that is, during time periods 1404c, 1404d and 1404e, whereby the position of the intruding object is corrected. In other words, the intruding object position correction processing using the inter-frame difference processing is executed at

10   the timing where the control operation of the camera universal head is not performed.

The operation of the third embodiment of the invention is explained with reference to Figs. 1A and 1B. Figs. 1A and 1B are flowcharts showing an example

15   of the processing according to the third embodiment of the invention. The processing shown in Fig. 1A is equivalent to the process of Fig. 3 with the search range enlarging loop (steps 111 to 114) added thereto. The other processing steps have already been explained

20   and are not explained again.

Fig. 1B is a flowchart showing the detail of the search range enlarging process (steps 112 to 114).

The search range enlarging loop (steps 111 to 114) is such that the search area 1003b providing the

25   projection range of the difference image due to the inter-frame difference shown in Fig. 8C is enlarged in steps (in five steps in this case). According to this embodiment, the search range enlarging process (steps

111 to 114) is repeated a plurality of times or, say, five times for each processing frame. In the second embodiment described above, the process is executed with the search range d as 50. According to this

5   embodiment, on the other hand, the initial value of the search range d is set to 10 (step 111a) in the first execution of the search range enlarging process including the search steps 111 to 114. Then, in the second search range enlarging process in step 114a,

10 d+10 is set as a new value d (d = 20). Similarly, in the third search range enlarging process, the value d is updated to 30 in step 114a, and then to 50 in the fifth search range enlarging process. In this way, the search range can be enlarged in steps. The search

15 range enlarging process is prevented by step 114 not to exceed five times. Steps 114a and 114b correspond to step 114 in Fig. 1A.

In the case where the difference of more than a predetermined value is detected in the difference

20 amount determining step 113 during the five sessions of executing steps 111 to 114, the repetition of the search range enlarging process is suspended and the intruding object position correcting step 115 is executed.

25   According to the second embodiment, in the case where a plurality of objects are moving in the intruding object search range, an intruding object is considered to exist in the portion of the largest

difference (as determined in the difference amount
determining step 113), and the position of the
intruding object is determined unconditionally. With
the processing according to the third embodiment, on
5   the other hand, a place near the position of the
intruding object detected by template matching where
the difference is large is determined as the position
of the intruding object. Even in the case where a
plurality of moving objects exist in the intruding
10   object search range, therefore, the position of the
intruding object can be correctly detected. This
effect is explained with reference to Fig. 10.

        Fig. 10 is a diagram for explaining a method
of detecting the position of an intruding object
15   according to the third embodiment of the invention, and
shows a difference image 1201 based on the inter-frame
difference at a given time point. Fig. 10 shows an
image displayed the moment an object 1203 comes to
appear in the monitor view field while the intruding
20   object 1202 is being tracked. A rectangle 1204
indicates the position of the intruding object detected
in the template matching step 106. Also, rectangles
1204a, 1204b, 1204c, 1204d and 1204e indicate the
search range in each step of enlarging the search
25   range.

        In the second embodiment of the invention,
the portion of the largest difference is detected in
the range of the rectangle 1204e. In the intruding

object position correcting step 115, therefore, the
position is corrected to a different object 1203 having
a larger difference amount, but not to the intruding
object being tracked. Thus, the different object 1203,

5 not the intruding object 1202 to be tracked, is tracked
in error. According to the third embodiment of the
invention, in contrast, the intruding object 1202 can
be grasped in a search range 1204b enlarged for the
second time, for example, from the intruding object

10 position 1204 detected by the template matching
process. Specifically, since the difference fails to
reach a predetermined value in the first search range
1204a, the search is conducted again in the next
enlarged search range 1204b. The difference reach the

15 predetermined value in this search range 1204b, and
therefore the position of the particular object 1202
can be correctly detected and corrected.

According to this embodiment, the number d of
pixels along X and Y axes indicating the search range

20 is set to 10 (d = 10) for the first search and 20 (d =
20) for the second search. In this way, the number of
pixels representing the search range is increased in
steps. Nevertheless, the method of setting the search
range is not limited. As another method of setting the

25 search range, the size of the search range is fixed to
the template size, and with respect to the position
1204 detected in the template matching step 106, the
search is conducted in the search range moved by 10

pixels upward, for example, followed by conducting the search in the search range moved upward by 10 pixels and leftward by 10 pixels, followed by conducting the search in the search range moved leftward by 10 pixels,

5 followed by conducting the search in the search range moved downward by 10 pixels and leftward by 10 pixels, followed by conducting the search in the search range moved downward by 10 pixels, followed by conducting the search in the search range moved downward by 10 pixels

10 and rightward by 10 pixels, followed by conducting the search in the search range moved rightward by 10 pixels, followed by conducting the search in the search range moved upward by 10 pixels and rightward by 10 pixels, followed by conducting the search in the search

15 range moved upward by 20 pixels, followed by conducting the search in the search range moved upward by 20 pixels and leftward by 20 pixels, followed by conducting the search in the search range moved leftward by 20 pixels, and so on. In this way, the

20 search is conducted while moving the search range spirally to detect a range associated with the maximum accumulated value of the difference or a range associated with accumulated value of the difference equal to or larger than a predetermined value.

25 Even in the case where a plurality of objects exist in a given intruding object search range, therefore, the correct position of an intruding object can be detected. As a result, the performance of the

intruding object tracking process is improved, and a reliable monitor system is constructed.

Although the embodiments have been explained as to the case where an intruding object is detected and tracked, an object to be detected is not limited to an intruding object but may be any moving member other than an intruding object. Further, although the embodiments have been explained as to the case where the invention is applied to a monitor system, the invention may be applied to a system for merely detecting and tracking a moving member without aiding the monitoring of the member. In this case, effects similar to that of the aforesaid embodiments can be obtained.

Further, in the present invention, the search range enlarging process of the third embodiment may be applied to the first embodiment.

This invention obviates the problem of the prior art that in the case where an intruding object changes in direction, the template may be displaced from the intruding object and come to contain many images other than the intruding object to be tracked, thereby making it impossible to hold the intruding object with the template matching the background portion undesirably. Further, according to this invention, even an intruding object changing considerably in direction can be detected and tracked. As a result, an intruding object tracking method and an

intruding object tracking apparatus are provided which
are high in reliability and stable in operation.

It should be further understood by those
skilled in the art that although the foregoing
5 description has been made on embodiments of the
invention, the invention is not limited thereto and
various changes and modifications may be made without
departing from the spirit of the invention and the
scope of the appended claims.